

# SEQUENCE LISTING

<110> Benning, Christoph  
Riekhof, Wayne  
Klug, Rouven

<120> Compositions and Methods for the Production of Betaine Lipids

<130> MSU-07769

<150> 10/118,495

<151> 2002-04-08

<160> 52

<170> PatentIn version 3.2

<210> 1

<211> 1252

<212> DNA

<213> Rhodobacter sphaeroides

<400> 1

gtgacgcagt tcgccctcac ccacctgccc gccccgccgg ttgcccccca gatcggcgcc	60
gccgtgcacc gcacgtcgct tctcagcgcc gaaggactga tggagcggat gttctcgcgc	120
ctcttcacg gcctcgtcta tccgcagatc tgggaggatc cggcgggtgga catggcggcc	180
ctcgccatcc gccccgggga cgggctggtg gccatcgctt cgggcgggtt caacgtgctt	240
tcctatctca cgcaggggccc gggctcgatc ctgcgcgtgg atctctcgcc cgcccatgtg	300
gcgctggggc ggctgaagct cgccgccgcg cggacgtgc ccgaccatgc cgccttcttc	360
gatctcttcg gtcgcgcaga cctgcccggc aatgcggccc tctacgaccg ccacatcgcg	420
cccgcgctcg acggccggag ccgccgctac tgggaggcgc gcagcccctt cggccggcgc	480
atccagctgt tcgagcgcg cttctaccgg caccgtgccc tcggccgctt catcggcgcg	540
gcccatagc tcgcgcgggc cgccgggcacc gacctgcggg gctttctcga ctgtcccac	600
atcgaggcgc agcgcagctt cttctacgcc catatcgggc cgctcttcga ggcgcccgtg	660
gtgcaggcgc tcgcccgacg gccggccgcg ctcttcgggc tggggatccc gcccgcgcaa	720
tatgcgcttc tggcgggaga cggcgacggc gacgtgctgc cggtgctgcg ccagcgccctc	780
caccggctgc tctgtgactt cccctgcgc gagaactact tcgccttcca ggccatcgcc	840
cgccgctatc cgcgggccgg cgagggcgcg ctgccgccct atctcgaacc caccgccttc	900
gagacgctgc gcgagaacgc gggccgggtg cagatcgaga accgcagcct gaccgaggcg	960
ctcgcgcccg aacccgagga gagcatccac ggcttcaccc tgctcgatgc gcaggactgg	1020
atgacggacg cgcagctgac cgcgctctgg cggcagggtga cgcgcactgc agcgccgggc	1080

gcgcgggtga tcttccgcac cggcggggcg gccgacctgc tgcccggccg agtgcccag	1140
gagatcctcg ggcactggcg cgccgaccgg gcggcgggac aggcgggcca tgccgccgac	1200
cgttcggcga tctacggcgg cttccacctc taccggcgga gggacgccat ga	1252

<210> 2  
 <211> 1255  
 <212> DNA  
 <213> Rhodobacter sphaeroides

<400> 2	
atgaccgacg ccacccatgc ggcgctgatg gacgcgacct accgccacca gcgcgggatc	60
tacgacgtca cgcggcgcca cttcctgctc ggccgcgacc ggctgatcgc cgagctcgac	120
ccgccccccg gcgcccgggt gctcgagatc gcctgcggca cggggcgcaa cctcgacctg	180
atcggccggc gctggcccgg ctgccggctc tcggggctcg acatctcgca ggagatgctg	240
gcctcgcccc gcgcgcgtct gggccggcgc gcgacgctgg cgctcgcgga tgccaccg	300
ttcgaggccc tgcccctctt cggcaccgac cggttcgagc ggatcgtcct ctctacgcg	360
ctctcgatga tccccgactg gcgcgaggcc ctgctgagg cggcgcttca tctcgtgccg	420
ggggggcggc tgcattgctg cgacttcggc gatcaggcgg gcctgcccgg ctggggccgc	480
gccggcctgc gcggctggat cgggcgcttc cacgtcacgc cgcgcgacga tctgggcacg	540
gcactgggcg aaacggcgct cgggatcggg ggctatgccg aataccggtc cctcggcggg	600
ggatatgcga ttctcggcac gctcacgcgg tgagagatcc cctgccctgc gcgtgacgt	660
tgtctgcccg caggcgaccg gccgcgcgac ggccggcctg cgggcgatcc ggcgactga	720
agggccggcg cgtcgcgcgc ggggacgtag cccgcagcgg caagcggccg acagagcctg	780
acagaccgtt cacggtgcgc gctccggatc ggggtgtggag ccggtgttgc agaggtcagg	840
cctcgaggga aagccctctg gcccgcggg caaattgtcc gggatctcta atcgggaaat	900
tggtcggagc gagaggattc gaacctccga cccctgctc ccgaagcagg tgcgctacca	960
ggctgcgcta cgctccgacc ttggcgtgcg gattataggg tcgcgcatcc gaatgcaagg	1020
gggtccgaac gcaattcgct acggagtgtc tcgcgtctcg cggcggcgca gaaggcgcgg	1080
catgaggccc acctcgggcc gcaggcgcgt ctggctcgcc gggcggttct ccgacacgtt	1140
gcggcgcgat tcgcggccga cgatatagag gccgctcgcg atgatgaccc ccgccccgac	1200
ccaggtccag acgtcggacc gctcgccgaa gatgagccag ccgaagatcc ctgac	1255

<210> 3  
 <211> 416  
 <212> PRT  
 <213> Rhodobacter sphaeroides

<400> 3

Met Thr Gln Phe Ala Leu Thr His Leu Pro Ala Pro Pro Val Ala Arg  
 1 5 10 15

Gln Ile Gly Ala Ala Val His Arg Thr Ser Leu Leu Ser Ala Glu Gly  
 20 25 30

Leu Met Glu Arg Met Phe Ser Arg Leu Phe His Gly Leu Val Tyr Pro  
 35 40 45

Gln Ile Trp Glu Asp Pro Ala Val Asp Met Ala Ala Leu Ala Ile Arg  
 50 55 60

Pro Gly Asp Arg Leu Val Ala Ile Ala Ser Gly Gly Cys Asn Val Leu  
 65 70 75 80

Ser Tyr Leu Thr Gln Gly Pro Gly Ser Ile Leu Ala Val Asp Leu Ser  
 85 90 95

Pro Ala His Val Ala Leu Gly Arg Leu Lys Leu Ala Ala Ala Arg Thr  
 100 105 110

Leu Pro Asp His Ala Ala Phe Phe Asp Leu Phe Gly Arg Ala Asp Leu  
 115 120 125

Pro Gly Asn Ala Ala Leu Tyr Asp Arg His Ile Ala Pro Ala Leu Asp  
 130 135 140

Gly Arg Ser Arg Arg Tyr Trp Glu Ala Arg Ser Pro Phe Gly Arg Arg  
 145 150 155 160

Ile Gln Leu Phe Glu Arg Gly Phe Tyr Arg His Gly Ala Leu Gly Arg  
 165 170 175

Phe Ile Gly Ala Ala His Thr Leu Ala Arg Ala Ala Gly Thr Asp Leu  
 180 185 190

Arg Gly Phe Leu Asp Cys Pro Asp Ile Glu Ala Gln Arg Ser Phe Phe  
 195 200 205

Tyr Ala His Ile Gly Pro Leu Phe Glu Ala Pro Val Val Gln Ala Leu  
 210 215 220

Ala Arg Arg Pro Ala Ala Leu Phe Gly Leu Gly Ile Pro Pro Ala Gln  
 225 230 235 240

Tyr Ala Leu Leu Ala Gly Asp Gly Asp Gly Asp Val Leu Pro Val Leu  
 245 250 255

Arg Gln Arg Leu His Arg Leu Leu Cys Asp Phe Pro Leu Arg Glu Asn  
 260 265 270

Tyr Phe Ala Phe Gln Ala Ile Ala Arg Arg Tyr Pro Arg Pro Gly Glu  
 275 280 285

Gly Ala Leu Pro Pro Tyr Leu Glu Pro Thr Ala Phe Glu Thr Leu Arg  
 290 295 300

Glu Asn Ala Gly Arg Val Gln Ile Glu Asn Arg Ser Leu Thr Glu Ala  
 305 310 315 320

Leu Ala Ala Glu Pro Glu Glu Ser Ile His Gly Phe Thr Leu Leu Asp  
 325 330 335

Ala Gln Asp Trp Met Thr Asp Ala Gln Leu Thr Ala Leu Trp Arg Gln  
 340 345 350

Val Thr Arg Thr Ala Ala Pro Gly Ala Arg Val Ile Phe Arg Thr Gly  
 355 360 365

Gly Ala Ala Asp Leu Leu Pro Gly Arg Val Pro Glu Glu Ile Leu Gly  
 370 375 380

His Trp Arg Ala Asp Arg Ala Ala Gly Gln Ala Gly His Ala Ala Asp  
 385 390 395 400

Arg Ser Ala Ile Tyr Gly Gly Phe His Leu Tyr Arg Arg Arg Asp Ala  
 405 410 415

<210> 4  
 <211> 210  
 <212> PRT  
 <213> Rhodobacter sphaeroides

<400> 4

Met Thr Asp Ala Thr His Ala Ala Leu Met Asp Ala Thr Tyr Arg His  
 1 5 10 15

Gln Arg Arg Ile Tyr Asp Val Thr Arg Arg His Phe Leu Leu Gly Arg  
 20 25 30

Asp Arg Leu Ile Ala Glu Leu Asp Pro Pro Pro Gly Ala Arg Val Leu  
 35 40 45

Glu Ile Ala Cys Gly Thr Gly Arg Asn Leu Asp Leu Ile Gly Arg Arg  
 50 55 60

Trp Pro Gly Cys Arg Leu Ser Gly Leu Asp Ile Ser Gln Glu Met Leu  
 65 70 75 80

Ala Ser Ala Arg Ala Arg Leu Gly Arg Arg Ala Thr Leu Ala Leu Gly  
 85 90 95

Asp Ala Thr Arg Phe Glu Ala Leu Pro Leu Phe Gly Thr Asp Arg Phe  
 100 105 110

Glu Arg Ile Val Leu Ser Tyr Ala Leu Ser Met Ile Pro Asp Trp Arg  
 115 120 125

Glu Ala Leu Arg Glu Ala Ala Leu His Leu Val Pro Gly Gly Arg Leu  
 130 135 140

His Val Val Asp Phe Gly Asp Gln Ala Gly Leu Pro Gly Trp Ala Arg  
 145 150 155 160

Ala Gly Leu Arg Gly Trp Ile Gly Arg Phe His Val Thr Pro Arg Asp  
 165 170 175

Asp Leu Gly Thr Ala Leu Gly Glu Thr Ala Leu Gly Ile Gly Gly Tyr  
 180 185 190

Ala Glu Tyr Arg Ser Leu Gly Gly Gly Tyr Ala Ile Leu Gly Thr Leu  
 195 200 205

Thr Arg  
 210

<210> 5  
 <211> 26  
 <212> DNA  
 <213> Rhodobacter sphaeroides

<400> 5  
 cgccctcacc cacattcccg ccccg

<210> 6  
 <211> 31  
 <212> DNA  
 <213> Rhodobacter sphaeroides  
  
 <400> 6  
 gactgtcccg agatcgaggg ccagcgccag c 31  
  
 <210> 7  
 <211> 26  
 <212> DNA  
 <213> Rhodobacter sphaeroides  
  
 <400> 7  
 gccgccgacc gtacggcgat ctacgg 26  
  
 <210> 8  
 <211> 28  
 <212> DNA  
 <213> Rhodobacter sphaeroides  
  
 <400> 8  
 gctgatggac gcgtcctacc gccaccag 28  
  
 <210> 9  
 <211> 29  
 <212> DNA  
 <213> Rhodobacter sphaeroides  
  
 <400> 9  
 cggttcgagc ggctcgctct ctctacgc 29  
  
 <210> 10  
 <211> 29  
 <212> DNA  
 <213> Rhodobacter sphaeroides  
  
 <400> 10  
 ggatatgcga ttctcgccac gctcacgcg 29  
  
 <210> 11  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Synthetic  
  
 <400> 11  
 gctctagaat ggcgagttc gccctc 26

<210> 12  
 <211> 30  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Synthetic  
  
 <400> 12  
 acatgcatgc aggacgatcc gctcgaaccg 30  
  
 <210> 13  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Synthetic  
  
 <400> 13  
 gctctagaat ggccgacgcc acccat 26  
  
 <210> 14  
 <211> 30  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Synthetic  
  
 <400> 14  
 acatgcatgc aggacgatcc gctcgaaccg 30  
  
 <210> 15  
 <211> 36  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Synthetic  
  
 <400> 15  
 atgcaggtgt ggcctccagt gacgcagttc gccctc 36  
  
 <210> 16  
 <211> 36  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Synthetic  
  
 <400> 16  
 gagggcgaac tgcgtcactg gaggccacac ctgcat 36

<210> 17  
 <211> 36  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Synthetic  
  
 <400> 17  
 atgcaggtgt ggcctccaat gaccgacgac acccat 36  
  
 <210> 18  
 <211> 36  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Synthetic  
  
 <400> 18  
 atgggtggcg tcggtcattg gaggccacac ctgcat 36  
  
 <210> 19  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Synthetic  
  
 <400> 19  
 gctctagaaa ccacaagaac taagaa 26  
  
 <210> 20  
 <211> 30  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Synthetic  
  
 <400> 20  
 acatgcatgc aggacgatcc gtcgaaccg 30  
  
 <210> 21  
 <211> 29  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Synthetic  
  
 <400> 21  
 acatgcatgc ctctaccgc gtgagcgtg 29



<210> 22  
 <211> 1248  
 <212> DNA  
 <213> Mesorhizobium loti

<400> 22  
 atgacggacg tctcctcgga tctgggtttt cgccgcggca aggaagttgg aaaggccgtc 60  
 taccagaacc gcgcgctttc caaagccggc atctccgagc ggctgttcgc cttcctgttt 120  
 tccggcctcg tctatccgca gatctgggaa gaccccgatg tcgacatgga ggccatgcag 180  
 cttgggtcagg gccatcgcat cgtcacaatc gcttccggcg gctgcaacat cctcgcctac 240  
 ctcacccgtt cgccggcacg gatcgacgcc gtcgacctca acgccgccca catcgcgctg 300  
 aaccgcatga agctggaggc ggtgcgccgt ctgccctcgc agggcgatct gttccgcttt 360  
 ttcggcgccg ccgacaccag ccacaattcg caagcctatg accgctttat tgcgccgcat 420  
 ctcgatccgg tcagccgccca ctattgggag cgccgcaact ggcgtggtcg ccggcgcatc 480  
 gccgtcttcg accgcaattt ctaccagacc ggcctgctcg gcctgttcat cgccatgggc 540  
 catcgcacgg cgaaattctt cggcgtcaac ccggcccaca tgatggaagc caggaatatc 600  
 ggcgagcagc gccgttctt caacgaggag ctggcgccgg tcttcgacaa gaagcttttg 660  
 aaatggggca cctcgcgtaa ggctcgcgtg ttcggcctcg gcattccgcc ggcgcgtagc 720  
 gattccctga tcacctcagg cgacggcacc atggccagcg ttctgaaggc ccggctggaa 780  
 aagctcgcct gcgattttcc cctggaaaac aattatttcg cctggcaggc ttttgcccgc 840  
 cgctatccaa atcccggta ggccgccttg cccgcctatc tggaaaagca gaactacgaa 900  
 accatccgcg gcaatatcga ccgcgtcgcc atccaccatg ccaatctgat cgaattcctc 960  
 gccggcaagg acgcgggcac cgtcgatcgc ttcactctgc tcgatgcgca ggactggatg 1020  
 accgatgacc agctcaacgc gctgtggtcg gaaatcagcc gcaccgcctc cgcaggcgcc 1080  
 cgcgtcatct tccgcaccgc cgccgagccc agcctgctgc caggccgcgt ctcgacctcg 1140  
 ctgctcgacc agtgggacta tcaggacgag gcgtcgcgcg aattctcggc acgcgaccgt 1200  
 tcggccatct atggcggctt ccacctctat gtgaagcgca cggcatga 1248

<210> 23  
 <211> 678  
 <212> DNA  
 <213> Mesorhizobium loti

<400> 23  
 atgaccgagc tgccggccag ccccgaattc aaggccaatc atgccgaact gatggacggc 60  
 gtctaccact ggcagcgcca catctatgac ctgactcgca aatactatct gctcggccgc 120  
 gaccggctga tcgatgggct tgaggtgccg caaggcgcca ccgtgctgga actcggctgc 180  
 ggcaccggcc gcaacatcat cctggccgcc cgccgctacc ctgatgcccg cttcttcggc 240

ctggatatct cggccgagat gctggagacg gccggcaagg cgatcgaccg cgaaggcctg	300
tccggccacg taacgctgac acgaggcgac gccaccgatt tcgacgccgc ggcactttac	360
ggcatcgagc gcttcgaccg cgtcttcgtc tcctattcgc tgtcgatgat cccaggctgg	420
gaaaagacgg tgcggcgcc actcgccgca ctatcccca acggctcgct gcacatcgtc	480
gatttcggcc agcaggaagg cctaccgggc tggttccgta ccttgctgcg cggttggctg	540
aaaaaattcc acgtaacgcc gcgtgaatcg ctgcgcgaag ttctggaatc ggaatctcgg	600
cgaaccggcg caaccttcg tttccgcacg ctttatcgcg gttacgcctg gctggcgatg	660
atcaagatcg ccagctaa	678

<210> 24  
 <211> 31  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic

<400> 24	
acatgcatgc aatgacggac gtctcctcgg a	31

<210> 25  
 <211> 29  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic

<400> 25	
cggggtacct catgccgtgc gcttcacat	29

<210> 26  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic

<400> 26	
gcggatccga tgaccgagct gccgg	25

<210> 27  
 <211> 28  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic

<400> 27  
 gcaagctttt agctggcgat cttgatca

28

<210> 28  
 <211> 1251  
 <212> DNA  
 <213> Agrobacterium tumefaciens

<400> 28  
 atgacgagtg cggcacccaa gaccggcttc agcaaaaaca cgaaactgaa gtccgcattg 60  
 ctccagcaca aggcactctc caaaagcggc ctgtccgaac ggttttttcgg cgtcctcttt 120  
 tccggcctcg tctatccgca gatctgggaa gaccccgaga tcgacatgga agcgatggag 180  
 cttggcgaag gccaccgcat cgtcaccatc ggctccggcg gctgcaacat gctggcctat 240  
 ctctcgcgca acccgccag catcgatgtg gtggacctca acccgacca catcgcgctg 300  
 aacaagctga agctcgctgc cttccgccat ctgcccggcc atcaggatgt ggtgcgccac 360  
 ttcggccgcg ccggcacccg cagcaacagc gtcggttatg accgtttcat cgccgagcat 420  
 ctggatgcca cgaccaaggc atactggctg aagcgcaccc tttccggccg ccgtcgattt 480  
 tcggtgttcg acaggaacat ctaccggacc ggctgctcg gccgtttcat cggcgccggc 540  
 cacatcatgg cccgcctgca cggcgtgaaa ctcaccgaaa tggccaagac ccggacgctg 600  
 gacgaacagc gccagttttt tgacagcaag gtcgcgccgc ttttcgacaa gccggtggtg 660  
 cgctggctga cgaagcgcaa gagctcgctt ttcggccttg gcattccgcc gcgccagtat 720  
 gacgagctgg caagcctttc cagcgacggc acggttgccct ccgtcctcaa ggagcggctg 780  
 gaaaagcttg cctgcaactt cccgctcagc gacaattatt tcgcctggca ggcttttgcg 840  
 cgccgttatc ccgagccgca tgagggtgcc ctgcccgtt atctcaagcc ggaatattac 900  
 gaaaagatcc gcaacaacac cgcgcgcgtc gcggtgcatc acgccaccta taccgagctg 960  
 ctttcccgca agccggcaaa tggcgtcgac cgctatatcc tgctcgatgc gcaggactgg 1020  
 atgacggatg tgcagctcaa cgagttatgg tcgcagatca gccgcactgc cgcacccggg 1080  
 gcacgcgtca tcttccgcac cgcggccgaa aagagcgcta tcgagggccg gctttcgccc 1140  
 gacatccgca accagtgggt ctatctcgaa gagcgctcca acgaactcaa cgccatggac 1200  
 cgctcggcca tttatggcgg cttccatatc taccagaggg ctatggcatg a 1251

<210> 29  
 <211> 416  
 <212> PRT  
 <213> Agrobacterium tumefaciens

<400> 29

Met	Thr	Ser	Ala	Ala	Pro	Lys	Thr	Gly	Phe	Ser	Lys	Asn	Thr	Lys	Leu	1	5	10	15
Lys	Ser	Ala	Leu	Leu	Gln	His	Lys	Ala	Leu	Ser	Lys	Ser	Gly	Leu	Ser	20	25	30	
Glu	Arg	Phe	Phe	Gly	Val	Leu	Phe	Ser	Gly	Leu	Val	Tyr	Pro	Gln	Ile	35	40	45	
Trp	Glu	Asp	Pro	Glu	Ile	Asp	Met	Glu	Ala	Met	Glu	Leu	Gly	Glu	Gly	50	55	60	
His	Arg	Ile	Val	Thr	Ile	Gly	Ser	Gly	Gly	Cys	Asn	Met	Leu	Ala	Tyr	65	70	75	80
Leu	Ser	Arg	Asn	Pro	Ala	Ser	Ile	Asp	Val	Val	Asp	Leu	Asn	Pro	His	85	90	95	
His	Ile	Ala	Leu	Asn	Lys	Leu	Lys	Leu	Ala	Ala	Phe	Arg	His	Leu	Pro	100	105	110	
Ala	His	Gln	Asp	Val	Val	Arg	His	Phe	Gly	Arg	Ala	Gly	Thr	Arg	Ser	115	120	125	
Asn	Ser	Val	Gly	Tyr	Asp	Arg	Phe	Ile	Ala	Glu	His	Leu	Asp	Ala	Thr	130	135	140	
Thr	Lys	Ala	Tyr	Trp	Ser	Lys	Arg	Thr	Leu	Ser	Gly	Arg	Arg	Arg	Ile	145	150	155	160
Ser	Val	Phe	Asp	Arg	Asn	Ile	Tyr	Arg	Thr	Gly	Leu	Leu	Gly	Arg	Phe	165	170	175	
Ile	Gly	Ala	Gly	His	Ile	Met	Ala	Arg	Leu	His	Gly	Val	Lys	Leu	Thr	180	185	190	
Glu	Met	Ala	Lys	Thr	Arg	Thr	Leu	Asp	Glu	Gln	Arg	Gln	Phe	Phe	Asp	195	200	205	
Ser	Lys	Val	Ala	Pro	Leu	Phe	Asp	Lys	Pro	Val	Val	Arg	Trp	Leu	Thr	210	215	220	

Lys Arg Lys Ser Ser Leu Phe Gly Leu Gly Ile Pro Pro Arg Gln Tyr  
 225 230 235 240

Asp Glu Leu Ala Ser Leu Ser Ser Asp Gly Thr Val Ala Ser Val Leu  
 245 250 255

Lys Glu Arg Leu Glu Lys Leu Ala Cys Asn Phe Pro Leu Ser Asp Asn  
 260 265 270

Tyr Phe Ala Trp Gln Ala Phe Ala Arg Arg Tyr Pro Glu Pro His Glu  
 275 280 285

Gly Ala Leu Pro Ala Tyr Leu Lys Pro Glu Tyr Tyr Glu Lys Ile Arg  
 290 295 300

Asn Asn Thr Ala Arg Val Ala Val His His Ala Thr Tyr Thr Glu Leu  
 305 310 315 320

Leu Ser Arg Lys Pro Ala Asn Gly Val Asp Arg Tyr Ile Leu Leu Asp  
 325 330 335

Ala Gln Asp Trp Met Thr Asp Val Gln Leu Asn Glu Leu Trp Ser Gln  
 340 345 350

Ile Ser Arg Thr Ala Ala Ser Gly Ala Arg Val Ile Phe Arg Thr Ala  
 355 360 365

Ala Glu Lys Ser Val Ile Glu Gly Arg Leu Ser Pro Asp Ile Arg Asn  
 370 375 380

Gln Trp Val Tyr Leu Glu Glu Arg Ser Asn Glu Leu Asn Ala Met Asp  
 385 390 395 400

Arg Ser Ala Ile Tyr Gly Gly Phe His Ile Tyr Gln Arg Ala Met Ala  
 405 410 415

<210> 30  
 <211> 720  
 <212> DNA  
 <213> Agrobacterium tumefaciens

<400> 30  
 atgaaaacca tcggcgagaa tgtcggcctt gcagacagcg cgcgatgcggg cttgatggac 60  
 cgcgatgtatc gccaccagcg ccatatctac gatatcaccc gcaaatatta tcttctgggc 120  
 cgtgaccgga ccatttccgg cctcgacgtg ccaaagggcg gcacgctgct ggaaatcggc 180  
 tgcggcaccg gccgcaacct gctgctggcc agccgccggt ttcccgcgcg caaactcttc 240

ggccctcgata tatcagccga aatgctgctg accgcctccg agaattttgc cggcaaagcg 300  
 gagcgaccca ttctgcgtgt cgccgatgcc accgctttcc ggtcttcgga attcggccag 360  
 cccgatggct tcgaccgct catgatccct tatgcgctgt cgatgatacc ggactgggaa 420  
 aaagcgatcg aacaggcgct cgcggcgctg aaacccggcg gttcgctgca tatcgctgat 480  
 ttcggccagc aggaacagtt gccgaagtgg ttccgcacgc ttcttcaagc ctggctcacc 540  
 cgctttcacg ttacgccccg cgcaaatttc cgttacgttc tcgccaatat ggccggccgt 600  
 ttcgacggga atctcgtctt cgaggaaatc gcgaggggat acgcatggcg ggctgtcatc 660  
 acgcttccgg ttgccgaagc cccgcagccg aagatccacc gcttattggc tgacgcctga 720

<210> 31  
 <211> 210  
 <212> PRT  
 <213> *Agrobacterium tumefaciens*

<400> 31

Met Thr Asp Ala Thr His Ala Ala Leu Met Asp Ala Thr Tyr Arg His  
1 5 10 15

Gln Arg Arg Ile Tyr Asp Val Thr Arg Arg His Phe Leu Leu Gly Arg  
20 25 30

Asp Arg Leu Ile Ala Glu Leu Asp Pro Pro Pro Gly Ala Arg Val Leu  
35 40 45

Glu Ile Ala Cys Gly Thr Gly Arg Asn Leu Asp Leu Ile Gly Arg Arg  
50 55 60

Trp Pro Gly Cys Arg Leu Ser Gly Leu Asp Ile Ser Gln Glu Met Leu  
65 70 75 80

Ala Ser Ala Arg Ala Arg Leu Gly Arg Arg Ala Thr Leu Ala Leu Gly  
85 90 95

Asp Ala Thr Arg Phe Glu Ala Leu Pro Leu Phe Gly Thr Asp Arg Phe  
100 105 110

Glu Arg Ile Val Leu Ser Tyr Ala Leu Ser Met Ile Pro Asp Trp Arg  
115 120 125

Glu Ala Leu Arg Glu Ala Ala Leu His Leu Val Pro Gly Gly Arg Leu  
130 135 140

His Val Val Asp Phe Gly Asp Gln Ala Gly Leu Pro Gly Trp Ala Arg  
145 150 155 160

Ala Gly Leu Arg Gly Trp Ile Gly Arg Phe His Val Thr Pro Arg Asp  
165 170 175

Asp Leu Gly Thr Ala Leu Gly Glu Thr Ala Leu Gly Ile Gly Gly Tyr  
180 185 190

Ala Glu Tyr Arg Ser Leu Gly Gly Tyr Ala Ile Leu Gly Thr Leu  
195 200 205

Thr Arg  
210

<210> 32  
<211> 1251  
<212> DNA  
<213> Sinorhizobium meliloti

<400> 32  
atgaccgact tcgccccgga tgccggcttc ggcaagaaga atccgaaact gaaaagcgca 60  
ctcctgcagc acaaagctct ctccccgcc ggtctctccg aacgcctgtt cgggctgctc 120  
ttttccggac tcgtctaccc gcagatctgg gaggaccgga ttgtcgacat ggaagcgatg 180  
cagatccgtc ccggacatcg gatcgtgacg atcggttccg gcggctgcaa catgctgacc 240  
tatctctccg ccgagcctgc ccggatagac gtggtcgatc tcaacccccca tcacatcgcg 300  
ctcaaccggc tgaagctgtc tgcccttctgc cacctgccga gccacaagga cgtgggtgcgg 360  
ttcctcgccg tcgaaggtac gcgcacgaat ggccaggcct acgacgtgtt cctcgcgccg 420  
aagctcgatc cggcaaccgc cgcctattgg aacggccgag atctcaccgg ccgccggcgc 480  
atcggcgtct tcgggcgcaa cgtttatcgt accggcctgc ttggccgttt catttccgcc 540  
agccatgctc tcgcacggct gcacggcatc aatccggaag atttcgtcaa ggcgcgctcc 600  
atgcgcgagc agcggcagtt ctctgacgac aagctcgctc cgctcttcga gcgtccggtc 660  
atccggttga tcaccagccg caagagctcc cttttcggcc tcggcatccc gccgcagcag 720  
ttcgacgaac tcgcgagcct gagccgggag aaatccgtcg ccgcggtgct gcgcaatcgc 780  
ctggaaaagc tgacctgtca tttccccttg cgcgataact acttcgcctg gcaggccttt 840  
gcacggcgct acccgcggcc ggacgagggc gagttgccac cttatcttca ggcacgcgga 900  
tacgaagcga ttcgcgacaa tgcggagcgc gtcgaggtcc accatgcgag cttcacggag 960  
cttctcgccg gcaagccgc cgcctcagtc gaccgctacg tgctcctcga cgcacaggac 1020  
tggatgaccg accagcagct gaacgacctc tggacggaga tcacccgcac cgccgacgcc 1080

ggcgcggtcg tgatcttccg cacggcggcc gaagcgagca tcctgccggg ggcctctccc 1140  
accaccctcc tcgatcagtg gtactatgat gccgagactt cgatgaggct cggcgctgaa 1200  
gaccggtcgg cgatctatgg cggcttcac atctaccgga agaaagcatg a 1251

<210> 33  
<211> 416  
<212> PRT  
<213> Sinorhizobium meliloti

<400> 33

Met Thr Asp Phe Ala Pro Asp Ala Gly Phe Gly Lys Lys Asn Pro Lys  
1 5 10 15

Leu Lys Ser Ala Leu Leu Gln His Lys Ala Leu Ser Pro Ala Gly Leu  
20 25 30

Ser Glu Arg Leu Phe Gly Leu Leu Phe Ser Gly Leu Val Tyr Pro Gln  
35 40 45

Ile Trp Glu Asp Pro Ile Val Asp Met Glu Ala Met Gln Ile Arg Pro  
50 55 60

Gly His Arg Ile Val Thr Ile Gly Ser Gly Gly Cys Asn Met Leu Thr  
65 70 75 80

Tyr Leu Ser Ala Glu Pro Ala Arg Ile Asp Val Val Asp Leu Asn Pro  
85 90 95

His His Ile Ala Leu Asn Arg Leu Lys Leu Ser Ala Phe Arg His Leu  
100 105 110

Pro Ser His Lys Asp Val Val Arg Phe Leu Ala Val Glu Gly Thr Arg  
115 120 125

Thr Asn Gly Gln Ala Tyr Asp Val Phe Leu Ala Pro Lys Leu Asp Pro  
130 135 140

Ala Thr Arg Ala Tyr Trp Asn Gly Arg Asp Leu Thr Gly Arg Arg Arg  
145 150 155 160

Ile Gly Val Phe Gly Arg Asn Val Tyr Arg Thr Gly Leu Leu Gly Arg  
165 170 175

Phe Ile Ser Ala Ser His Ala Leu Ala Arg Leu His Gly Ile Asn Pro  
180 185 190



Glu	Asp	Phe	Val	Lys	Ala	Arg	Ser	Met	Arg	Glu	Gln	Arg	Gln	Phe	Phe	195	200	205
Asp	Asp	Lys	Leu	Ala	Pro	Leu	Phe	Glu	Arg	Pro	Val	Ile	Arg	Trp	Ile	210	215	220
Thr	Ser	Arg	Lys	Ser	Ser	Leu	Phe	Gly	Leu	Gly	Ile	Pro	Pro	Gln	Gln	225	230	235
Phe	Asp	Glu	Leu	Ala	Ser	Leu	Ser	Arg	Glu	Lys	Ser	Val	Ala	Ala	Val	245	250	255
Leu	Arg	Asn	Arg	Leu	Glu	Lys	Leu	Thr	Cys	His	Phe	Pro	Leu	Arg	Asp	260	265	270
Asn	Tyr	Phe	Ala	Trp	Gln	Ala	Phe	Ala	Arg	Arg	Tyr	Pro	Arg	Pro	Asp	275	280	285
Glu	Gly	Glu	Leu	Pro	Pro	Tyr	Leu	Gln	Ala	Ser	Arg	Tyr	Glu	Ala	Ile	290	295	300
Arg	Asp	Asn	Ala	Glu	Arg	Val	Glu	Val	His	His	Ala	Ser	Phe	Thr	Glu	305	310	315
Leu	Leu	Ala	Gly	Lys	Pro	Ala	Ala	Ser	Val	Asp	Arg	Tyr	Val	Leu	Leu	325	330	335
Asp	Ala	Gln	Asp	Trp	Met	Thr	Asp	Gln	Gln	Leu	Asn	Asp	Leu	Trp	Thr	340	345	350
Glu	Ile	Thr	Arg	Thr	Ala	Asp	Ala	Gly	Ala	Val	Val	Ile	Phe	Arg	Thr	355	360	365
Ala	Ala	Glu	Ala	Ser	Ile	Leu	Pro	Gly	Arg	Leu	Ser	Thr	Thr	Leu	Leu	370	375	380
Asp	Gln	Trp	Tyr	Tyr	Asp	Ala	Glu	Thr	Ser	Met	Arg	Leu	Gly	Ala	Glu	385	390	395
Asp	Arg	Ser	Ala	Ile	Tyr	Gly	Gly	Phe	His	Ile	Tyr	Arg	Lys	Lys	Ala	405	410	415

<210> 34  
 <211> 666  
 <212> DNA  
 <213> Sinorhizobium meliloti

<400> 34  
 atgagcgccg tgcagaccgc gaatgaaagc cacgctcatc tgatggaccg catgtatcgc 60  
 taccagcggg acatctatga ttctactcgc aaatactatc tcttcggccg tgacacgctg 120  
 atccgtgaac tgaacccgcc gccaggcgca tcgggtgctg aagtcggctg cggcacgggc 180  
 cgcaatctcg ccgtgatcgg ggatctctac cccgggtgcgc gcctcttcgg cctcgatatc 240  
 tcggccgaaa tgctggcgac cgccaaagcc aagctccggc gccaaaatcg gccggacgca 300  
 gtgttgccgg tcgccgacgc gacgaatttc accgccgcct cattcgatca ggaaggcttc 360  
 gaccggatcg tcatttccta cgccctttcc atgggtcccg aatgggaaaa ggcggtcgat 420  
 gccgcgattg ccgcgctcaa gccgggcggc tcgctgcata tcgccgactt cggccagcag 480  
 gaaggttggc cggccggctt ccgccgcttc ctccaggcct ggctcagacg cttccacgtc 540  
 acgccgcgcg aaacgctttt cgatgtgatg cgcaaaagag ccgagagaaa cggagcggcg 600  
 ctcgaggtca gatcgctgag acgaggttat gcctggcttg tcgtctatcg ccgcgcggca 660  
 ccgtag 666

<210> 35  
 <211> 221  
 <212> PRT  
 <213> Sinorhizobium meliloti

<400> 35

Met	Ser	Ala	Val	Gln	Thr	Ala	Asn	Glu	Ser	His	Ala	His	Leu	Met	Asp
1				5					10					15	
Arg	Met	Tyr	Arg	Tyr	Gln	Arg	Tyr	Ile	Tyr	Asp	Phe	Thr	Arg	Lys	Tyr
			20					25					30		
Tyr	Leu	Phe	Gly	Arg	Asp	Thr	Leu	Ile	Arg	Glu	Leu	Asn	Pro	Pro	Pro
		35					40					45			
Gly	Ala	Ser	Val	Leu	Glu	Val	Gly	Cys	Gly	Thr	Gly	Arg	Asn	Leu	Ala
	50					55					60				
Val	Ile	Gly	Asp	Leu	Tyr	Pro	Gly	Ala	Arg	Leu	Phe	Gly	Leu	Asp	Ile
65					70					75				80	
Ser	Ala	Glu	Met	Leu	Ala	Thr	Ala	Lys	Ala	Lys	Leu	Arg	Arg	Gln	Asn
				85					90					95	

Arg Pro Asp Ala Val Leu Arg Val Ala Asp Ala Thr Asn Phe Thr Ala  
100 105 110

Ala Ser Phe Asp Gln Glu Gly Phe Asp Arg Ile Val Ile Ser Tyr Ala  
115 120 125

Leu Ser Met Val Pro Glu Trp Glu Lys Ala Val Asp Ala Ala Ile Ala  
130 135 140

Ala Leu Lys Pro Gly Gly Ser Leu His Ile Ala Asp Phe Gly Gln Gln  
145 150 155 160

Glu Gly Trp Pro Ala Gly Phe Arg Arg Phe Leu Gln Ala Trp Leu Arg  
165 170 175

Arg Phe His Val Thr Pro Arg Glu Thr Leu Phe Asp Val Met Arg Lys  
180 185 190

Arg Ala Glu Arg Asn Gly Ala Ala Leu Glu Val Arg Ser Leu Arg Arg  
195 200 205

Gly Tyr Ala Trp Leu Val Val Tyr Arg Arg Ala Ala Pro  
210 215 220

<210> 36  
<211> 29  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Synthetic

<400> 36  
acatgcatgc agtgacgcag ttcgccctc

29

<210> 37  
<211> 29  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Synthetic

<400> 37  
cggggtacca ggacgatccg ctcgaaccg

29

<210> 38  
 <211> 89  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> 39  
  
 <400> 38  
 atgagaggat cgcacaccca tcacccatcac ggatccgcat gcgagctcgg taccgccgggt 60  
 cgacctgcag ccaagcttaa ttagctgag 89  
  
 <210> 39  
 <211> 91  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Synthetic  
  
 <400> 39  
 atgagaggat ctcatcacca tcacccatcac acggatccgc atgcgagctc ggtaccgccgg 60  
 gtcgacctgc agccaagctt aattagctga g 91  
  
 <210> 40  
 <211> 90  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Synthetic  
  
 <400> 40  
 atgagaggat ctcatcacca tcacccatcac gggatccgca tgcgagctcg gtaccgccggg 60  
 tcgacctgca gccaaagctta attagctgag 90  
  
 <210> 41  
 <211> 415  
 <212> PRT  
 <213> Mesorhizobium loti  
  
 <400> 41  
 Met Thr Asp Val Ser Ser Asp Leu Val Phe Arg Arg Gly Lys Glu Val  
 1 5 10 15  
  
 Gly Lys Ala Val Tyr Gln Asn Arg Ala Leu Ser Lys Ala Gly Ile Ser  
 20 25 30  
  
 Glu Arg Leu Phe Ala Phe Leu Phe Ser Gly Leu Val Tyr Pro Gln Ile  
 35 40 45

Trp Glu Asp Pro Asp Val Asp Met Glu Ala Met Gln Leu Gly Gln Gly  
 50 55 60

His Arg Ile Val Thr Ile Ala Ser Gly Gly Cys Asn Ile Leu Ala Tyr  
 65 70 75 80

Leu Thr Arg Ser Pro Ala Arg Ile Asp Ala Val Asp Leu Asn Ala Ala  
 85 90 95

His Ile Ala Leu Asn Arg Met Lys Leu Glu Ala Val Arg Arg Leu Pro  
 100 105 110

Ser Gln Gly Asp Leu Phe Arg Phe Phe Gly Ala Ala Asp Thr Ser His  
 115 120 125

Asn Ser Gln Ala Tyr Asp Arg Phe Ile Ala Pro His Leu Asp Pro Val  
 130 135 140

Ser Arg His Tyr Trp Glu Arg Arg Asn Trp Arg Gly Arg Arg Arg Ile  
 145 150 155 160

Ala Val Phe Asp Arg Asn Phe Tyr Gln Thr Gly Leu Leu Gly Leu Phe  
 165 170 175

Ile Ala Met Gly His Arg Thr Ala Lys Phe Phe Gly Val Asn Pro Ala  
 180 185 190

His Met Met Glu Ala Arg Asn Ile Gly Glu Gln Arg Arg Phe Phe Asn  
 195 200 205

Glu Glu Leu Ala Pro Val Phe Asp Lys Lys Leu Leu Lys Trp Ala Thr  
 210 215 220

Ser Arg Lys Ala Ser Leu Phe Gly Leu Gly Ile Pro Pro Ala Gln Tyr  
 225 230 235 240

Asp Ser Leu Ile Thr Ser Gly Asp Gly Thr Met Ala Ser Val Leu Lys  
 245 250 255

Ala Arg Leu Glu Lys Leu Ala Cys Asp Phe Pro Leu Glu Asn Asn Tyr  
 260 265 270

Phe Ala Trp Gln Ala Phe Ala Arg Arg Tyr Pro Asn Pro Gly Glu Ala  
 275 280 285

Ala Leu Pro Ala Tyr Leu Glu Lys Gln Asn Tyr Glu Thr Ile Arg Gly  
 290 295 300

Asn Ile Asp Arg Val Ala Ile His His Ala Asn Leu Ile Glu Phe Leu  
 305 310 315 320

Ala Gly Lys Asp Ala Gly Thr Val Asp Arg Phe Ile Leu Leu Asp Ala  
 325 330 335

Gln Asp Trp Met Thr Asp Asp Gln Leu Asn Ala Leu Trp Ser Glu Ile  
 340 345 350

Ser Arg Thr Ala Ser Ala Gly Ala Arg Val Ile Phe Arg Thr Ala Ala  
 355 360 365

Glu Pro Ser Leu Leu Pro Gly Arg Val Ser Thr Ser Leu Leu Asp Gln  
 370 375 380

Trp Asp Tyr Gln Asp Glu Ala Ser Arg Glu Phe Ser Ala Arg Asp Arg  
 385 390 395 400

Ser Ala Ile Tyr Gly Gly Phe His Leu Tyr Val Lys Arg Thr Ala  
 405 410 415

<210> 42  
 <211> 225  
 <212> PRT  
 <213> Mesorhizobium loti  
 <400> 42

Met Thr Glu Leu Pro Ala Ser Pro Glu Phe Lys Ala Asn His Ala Glu  
 1 5 10 15

Leu Met Asp Gly Val Tyr His Trp Gln Arg His Ile Tyr Asp Leu Thr  
 20 25 30

Arg Lys Tyr Tyr Leu Leu Gly Arg Asp Arg Leu Ile Asp Gly Leu Glu  
 35 40 45

Val Pro Gln Gly Gly Thr Val Leu Glu Leu Gly Cys Gly Thr Gly Arg  
 50 55 60

Asn Ile Ile Leu Ala Ala Arg Arg Tyr Pro Asp Ala Arg Phe Phe Gly  
 65 70 75 80

Leu Asp Ile Ser Ala Glu Met Leu Glu Thr Ala Gly Lys Ala Ile Asp  
 85 90 95

Arg Glu Gly Leu Ser Gly His Val Thr Leu Thr Arg Gly Asp Ala Thr  
 100 105 110

Asp Phe Asp Ala Ala Ala Leu Tyr Gly Ile Glu Arg Phe Asp Arg Val  
115 120 125

Phe Val Ser Tyr Ser Leu Ser Met Ile Pro Gly Trp Glu Lys Thr Val  
130 135 140

Ser Ala Ala Leu Ala Ala Leu Ser Pro Asn Gly Ser Leu His Ile Val  
145 150 155 160

Asp Phe Gly Gln Gln Glu Gly Leu Pro Gly Trp Phe Arg Thr Leu Leu  
165 170 175

Arg Gly Trp Leu Lys Lys Phe His Val Thr Pro Arg Glu Ser Leu Arg  
180 185 190

Glu Val Leu Glu Ser Glu Ser Arg Arg Thr Gly Ala Thr Phe Arg Phe  
195 200 205

Arg Thr Leu Tyr Arg Gly Tyr Ala Trp Leu Ala Met Ile Lys Ile Ala  
210 215 220

Ser  
225

<210> 43  
<211> 5242  
<212> DNA  
<213> Chlamydomonas reinhardtii

<400> 43  
tgttttattgc cgagcgggtgc gcggcgtgac cgccgtgggg ccccagggcc gggcactgcc 60  
tgtcacggag gctggtgccc tggcgggaca ccaacttgca actccggaac actcgcgcgg 120  
ggtgtgaccg cgaattgagg tctatatgaa agtataattg ctccggttga cgagacaccc 180  
tgccgcgact gatggagtga tggcgccaaa ttgacgcttt attctgcgat agttgcaatg 240  
gcgcagtctc cagcgccata tgttgacatg ctgtattaat tgcgactag cgctagcgct 300  
gcgcggcgcg gcattgggggt cacctgcgtg tccggtgtcc catcgaccat cgtgcggaat 360  
gagcccgcga cgtgctcata atagtttggt tcccgcgcat gctctttgtc gcggccttgt 420  
gcgtgcagac ctacatcttt gtgacactca acaaccgtca ttaaaccaaa gccccttaag 480  
tgctgttgat tagcggaacc atggggtcgg gtcgtgacgg ccggcctgcg agctacacca 540  
agaagaactt ctccctggag aagctcaagc tcagcagcat gaaggatgac ctgaccgttc 600  
tgcgccatat gtggttcggc agcaagaagg gcgatgatca cgctgctcgc ctggagagct 660  
tctacggggc ccaggccgct gcctgtaagt cgcgattaac tttatgctca ttttaagtgtt 720

tcgagttgct	tacgcgctgt	ctgcccgcgc	agatgatgct	ttccggtcgc	ggttcctctg	780
gggtcgcagg	cccatgctcg	ctgcagttgc	tgccgcgctg	gccgagcgc	cgaacctcat	840
ctgggttgac	ctgggtggtg	gcactggggg	gagtcgcggt	tgtatcgggg	aggtgcggga	900
tagccctggc	cgtttggtg	cgcctcgcgc	gccatttatg	ctgatctgcg	ctttgtcttt	960
gtgcgcgcag	gagaatgtcg	atatgatggc	tgattacatc	gacctggcga	agttcaagtc	1020
catctacgtg	gtcgacctgt	gccactcgct	gtgcgaggtg	gccaagaaga	aggcgaaggc	1080
caagggctgg	aagaatgtcc	aggctcgtga	ggccgacgct	tgccaatttg	cgccccctga	1140
gggcaccgcg	acgtcatca	ccttctccta	ctcgtcacg	agtgcggtg	aacgccgtcg	1200
acttgccatc	ggaggatcca	tccaccaca	acgttcatcc	cctctcacc	cgcgcttttt	1260
gctgttgacg	tgattccacc	gttccacaac	gtcatcgacc	aggcttgctc	gtacctgtcc	1320
caagacggcc	tggtgggctg	tgccgacttc	tacgtgagcg	gcaagtacga	cctgccccctg	1380
cgccagatgc	cctggtcgcg	ccgtttcttc	tgccggtgag	ttaccaggcc	gcggctacac	1440
atctctcaag	caagcgtagc	tgagcggggc	acagaaggcg	cttagggccc	ggggtgcgct	1500
tctgggtagt	ggcaaggata	tcggcaaggc	ccttcagcta	cccgaggtag	catggtgatc	1560
tgtcacgcag	ctcatcacgg	tgctgactga	ggcacgtgct	cgagtctgcg	gtcgtcatag	1620
ggctttcagc	acatcacagc	gcagtctccg	aatgcgcgtg	ctgcagcgtg	tatgctccgg	1680
gacagcacca	tgcaccgtgc	ggcgttcttg	acttgcaccc	gctgctctcg	ggcctcccct	1740
gggttcgctc	agacacgtct	ccgctgcccc	tgtgtctgct	gcgtcgtcgt	gtgcagatcg	1800
atcttcgaca	tcgacaacat	tgacatcggc	cccagcgcgc	gcgcctacct	ggagcagaag	1860
ctggagcgcg	tgtgggagca	gaacaccag	gtgtgtgccg	cggtgacacc	tggcctgtgc	1920
aagcccaagt	aaggtttgct	catcagccct	gcagcaccgc	cggcgtgcgg	cggatattca	1980
tcggcagagc	cgtccgcact	gttggttaat	ccaccagtcc	gcctgccttt	cctctgcgtg	2040
tttcagggtt	cgatccccta	cgtgccgtgg	ctgcgcgccc	cctactacgt	gtggattggc	2100
cgctgcccc	gcgttgggcg	tgagtcgcgt	cgcgcgggac	ctcgtccgca	tcaggacaca	2160
tggttacggg	ggcctcagtg	aagcgtagtc	tatcgttgca	gtcgtactgt	caagcgtagc	2220
aacaagcgcg	tccttgccgc	gttgaaattg	ttgggcgcaa	gcatgagccg	tgccgccgat	2280
gccccactcg	cgttgacct	gcagacgccc	tgcacgagga	gcgcgtggag	cggccgcccc	2340
tgttcccgcc	caccttcctg	tacacgcagt	cgtgggagga	ccccgagccg	gatattggagg	2400
tgagcggctc	agcctgaggc	agatacgcac	gaaggaaacg	ctcgtggtac	cccagggggc	2460
ggatcacgca	cgttcggaag	cattgtggca	cggtcacatc	gtcgcaccag	caaagcaagc	2520
actaagcacc	agcggttctg	tgggcaatgt	gccaccgcaa	tgccctggcca	gtgcgcagct	2580



tcgcattagt	gtatgccagt	atcacgccta	gctcagcctg	caagctgctg	tcgtagaaaag	2640
cagccgatgg	tggcactgag	cgcgaacact	cctgctgaaa	ctgtgcctgc	catcctctcg	2700
tgcccttccc	aggtgatgga	gatcaacccc	aaggacacgg	tgttgaccct	gactagcggc	2760
ggctgcaatg	ccctgaacct	gctggtgcag	ggggccggcc	aggtgagatc	atagtgcgtg	2820
ctgcaccctt	tctggtgtcc	ttacttcatg	atggccgctg	cacggaatgt	taggaagcgg	2880
ttggcataca	tgtagctgca	gaatgtcgct	gacgtcgtct	tccgcgctgc	tgtgcaggt	2940
ggtgtcggtg	gactgcaacc	ccgcgcagtc	ggcgcttctg	gagctgaaga	aggtggccat	3000
tcagcagctg	gagtttgagg	acgtgtggca	ggtaaggggc	tcctcatctg	gcgcctaggg	3060
actgcacacg	tctgttccta	ctgtgatgca	acatgcggga	cctggtcagg	ccttgccgca	3120
ttcgcaactg	gatgctgtcc	cccagggctc	agcacatcct	ctcgccggct	gcaacggcac	3180
agcccttttt	acactccgca	acccctgtct	tggtaattcc	ccttccacag	ctgttcggcg	3240
agggcgtgca	cccgcgcatt	gaggagctgt	acgagaagaa	gctggcgccc	ttcctgtcgc	3300
aaaccagcca	caacttctgg	tccaagcgcc	tctggtactt	ccagcacggc	ctgtactacc	3360
agggcgccat	gggcaagctg	tgttggtg	tgcagtgcct	ggccgtgggtg	ctgggactgg	3420
gcaagaccgt	caagcgcctc	gccaacgcgc	ccacaatgga	ggagcagcgc	cgtctgtggg	3480
acagcaacat	gctcatccac	ttcgtgaaga	acgggcccac	gccgctgggtg	tggctgttcg	3540
tcaagtctgt	gagcctgggtg	ctcttcaaca	aggccgtgct	gtgggttcggc	ggcggcgtgc	3600
cgggcaagca	gtacgcgctg	atcaaggcgg	acggcatccc	cattgagaac	tacatcgcgc	3660
gcaccatgga	cggcgtggcg	gagaactcgc	acgtgcgcaa	gcagaactac	ttctactaca	3720
actgcctcac	cggcaagtcc	ctgcgcgaca	actgccccac	ctacctgcgc	gaggcggcct	3780
tcgccaccct	caagagtggc	gtggtggaca	acctgaccgt	ctccaccaac	ttcttcatgg	3840
aggagctcaa	agcgcgcacc	tacaccaagg	tgggagccag	agcggagagt	gtgaaaagcg	3900
ggagatggcg	ggcgttgtag	gtgcatggcg	aacgcttgct	tgccaggggtg	aaagggcaga	3960
gcgagggaaa	cgggaatttc	gaattcggat	ggctgggaag	cggaaacgaa	caagaacgga	4020
atagaaggcc	gctggcgaga	aactggtgcg	ttggcgtttg	gcatccagtc	tgccaggcgc	4080
ccccgtgcct	acctgctcag	gtccgtctaa	ccgttacaac	cctcgccctc	accgcccccc	4140
ccccgcccc	ccaggtgatt	ctgatggacc	acgtggactg	gctggatatg	cccgtggcca	4200
acgagctggc	cgagtgcctg	gccaagcagg	ttgcgccggg	cggcatcgtc	atctggcgct	4260
ccgcctccct	cagcccgccc	tacgccgagc	tgatccagaa	ggcgggcttc	gacgtgcgct	4320
gcatccgccg	cgcactcag	ggctacatgg	accgcgtcaa	catgtacagg	tgcgtgcttt	4380
gtccgtgcag	gtgtgaggat	gtaggatgtc	agtggtgccg	ggagggagca	tgcgggtcga	4440

atgggaggag gttgcgttat tgatgcccga tgggacctgg gtgaatgcgg tgagcggcta	4500
ccgtatgtgc caaggacggg gttggcggcg tcgggcggca gcagaaactt gaaaggttgt	4560
tttagaagat caagatatgt tgcgataacg gaagcgacct agcgtcaggg cgcaacagcg	4620
cgtgggcacc gccgaaatct gcgttgctct tcgatgtata ccctgatcca atcttgctcg	4680
tttctccaat gcgcagctcc ttctacatgg ccgcccggaa gggcgccaag aaggacaact	4740
aagcagctgg cggcgaaggc acggcgggca agtggccggc tagcaacagc caatggcgct	4800
gacatcgcag gagcagtgcg ttggttgcta gccggcggcg tgccgtgcag ggagcaccgc	4860
tgtggtcaag ctgcggggac tgggctggca gcggaagcag tacaggcagc ggcgcaatgg	4920
gcccgttact cgtgcggggc ctgcagtcca gctcgaatca ttgcagcttg ctagcccggc	4980
tgtcacagca gcgcttcggt aggtgcggcc gtccacgttt agcgtgctgg actccatagt	5040
agtggtgcca gtaggagtta ggacaatgac cagtagcgca gttgaggacg ggcattttcc	5100
cggtagccca ggttccgctg tctgtggttg gtgtgtggca atccacgccg gacgcagttg	5160
taggggacac cgggtcgccg cacaggcttc ctggccactg ctgtgtgggg tgtgggggac	5220
cgatagcgcc cgaggcagca cc	5242

<210> 44  
 <211> 1947  
 <212> DNA  
 <213> Chlamydomonas reinhardtii

<400> 44	
atggggtcgg gtcgtgacgg ccggcctgcg agctacacca agaagaactt ctccctggag	60
aagctcaagc tcagcagcat gaaggatgac ctgaccgttc tgcgccatat gtggttcggc	120
agcaagaagg gcgatgatca cgctgctcgc ctggagagct tctacggggc ccaggccgct	180
gcctttgctg ccgcctggc cgagcgctcg aacctcatct gggttgacct ggggtggtggc	240
actggggaga atgtcgatat gatggctgat tacatcgacc tggcgaagtt caagtccatc	300
tacgtggctg acctgtgcca ctcgctgtgc gaggtggcca agaagaaggc gaaggccaag	360
ggctggaaga atgtccaggt cgtggaggcc gacgcttgcc aatttgcgcc ccctgagggc	420
accgcgacgc tcatcacctt ctctactcg ctacgatga ttccaccgtt ccacaacgtc	480
atcgaccagg cttgctcgta cctgtcccaa gacggcctgg tgggcgttg cgcacttctac	540
gtgagcggca agtacgacct gccctgcgc cagatgccct ggtcgcgccg tttcttcttg	600
cgatcgatct tcgacatcga caacattgac atcgcccccg agcgccgcgc ctacctggag	660
cagaagctgg agcgcgtgtg ggagcagaac acccagggtt cgatccccta cgtgccgtgg	720
ctgcgcgcc cctactacgt gtggattggc cgctgcca gcgttgcca cgcctgcac	780
gaggagcgcg tggagcggcc gcccatgttc ccgccacct tcctgtacac gcagtcgtgg	840

gaggaccccg agccggatat ggaggtgatg gagatcaacc ccaaggacac ggtgctgacc 900  
ctgactagcg gcggctgcaa tgcctgaac ctgctggtgc agggggccgg ccaggtggtg 960  
tcggtggact gcaaccccg gcagtcggcg cttctggagc tgaagaaggt ggccattcag 1020  
cagctggagt ttgaggacgt gtggcagctg ttcggcgagg gcgtgcaccc gcgcattgag 1080  
gagctgtacg agaagaagct ggcgcccttc ctgtcgcaaa ccagccacaa cttctggtcc 1140  
aagcgctctt ggtacttcca gcacggcctg tactaccagg gcggcatggg caagctgtgc 1200  
tgggtgctgc agtgcctggc cgtggtgctg ggactgggca agaccgtcaa gcgcctcgcc 1260  
aacgcgcca caatggagga gcagcgccgt ctgtgggaca gcaacatgct catccacttc 1320  
gtgaagaacg ggcccaagcc gctggtgtgg ctgttcgtca agttcgtgag cctggtgctc 1380  
ttcaacaagg ccgtgctgtg gttcggcggc ggcgtgccgg gcaagcagta cgcgtgatc 1440  
aaggcggacg gcatcccat tgagaactac atcgcgcgca ccattggacgg cgtggcggag 1500  
aactcgacg tgcgcaagca gaactacttc tactacaact gcctcaccgg caagttcctg 1560  
cgcgacaact gcccaccta cctgcgcgag gcggccttcg ccaccctcaa gaggggcgtg 1620  
gtggacaacc tgaccgtctc caccaacttc ttcattggagg agtcaaagc gcgcacctac 1680  
accaaggtga ttctgatgga ccacgtggac tggctggata tgcccggtggc caacgagctg 1740  
gccgagtgcc tggccaagca ggttgcgccg ggcggcatcg tcatctggcg ctccgcctcc 1800  
ctcagcccg cctacgccga gctgatccag aaggcgggct tcgacgtgcg ctgcatccgc 1860  
cgcgccactc agggctacat ggaccgcgtc aacatgtaca gtccttcta catggcccg 1920  
cggaagggcg ccaagaagga caactaa 1947

<210> 45  
<211> 648  
<212> PRT  
<213> Chlamydomonas reinhardtii

<400> 45

Met Gly Ser Gly Arg Asp Gly Arg Pro Ala Ser Tyr Thr Lys Lys Asn  
1 5 10 15

Phe Ser Leu Glu Lys Leu Lys Leu Ser Ser Met Lys Asp Asp Leu Thr  
20 25 30

Val Leu Arg His Met Trp Phe Gly Ser Lys Lys Gly Asp Asp His Ala  
35 40 45

Ala Arg Leu Glu Ser Phe Tyr Gly Pro Gln Ala Ala Ala Phe Ala Ala  
50 55 60

Arg Leu Ala Glu Arg Ser Asn Leu Ile Trp Val Asp Leu Gly Gly Gly  
 65 70 75 80  
 Thr Gly Glu Asn Val Asp Met Met Ala Asp Tyr Ile Asp Leu Ala Lys  
 85 90 95  
 Phe Lys Ser Ile Tyr Val Val Asp Leu Cys His Ser Leu Cys Glu Val  
 100 105 110  
 Ala Lys Lys Lys Ala Lys Ala Lys Gly Trp Lys Asn Val Gln Val Val  
 115 120 125  
 Glu Ala Asp Ala Cys Gln Phe Ala Pro Pro Glu Gly Thr Ala Thr Leu  
 130 135 140  
 Ile Thr Phe Ser Tyr Ser Leu Thr Met Ile Pro Pro Phe His Asn Val  
 145 150 155 160  
 Ile Asp Gln Ala Cys Ser Tyr Leu Ser Gln Asp Gly Leu Val Gly Val  
 165 170 175  
 Ala Asp Phe Tyr Val Ser Gly Lys Tyr Asp Leu Pro Leu Arg Gln Met  
 180 185 190  
 Pro Trp Ser Arg Arg Phe Phe Trp Arg Ser Ile Phe Asp Ile Asp Asn  
 195 200 205  
 Ile Asp Ile Gly Pro Glu Arg Arg Ala Tyr Leu Glu Gln Lys Leu Glu  
 210 215 220  
 Arg Val Trp Glu Gln Asn Thr Gln Gly Ser Ile Pro Tyr Val Pro Trp  
 225 230 235 240  
 Leu Arg Ala Pro Tyr Tyr Val Trp Ile Gly Arg Leu Pro Ser Val Gly  
 245 250 255  
 His Ala Leu His Glu Glu Arg Val Glu Arg Pro Pro Met Phe Pro Pro  
 260 265 270  
 Thr Phe Leu Tyr Thr Gln Ser Trp Glu Asp Pro Glu Pro Asp Met Glu  
 275 280 285  
 Val Met Glu Ile Asn Pro Lys Asp Thr Val Leu Thr Leu Thr Ser Gly  
 290 295 300  
 Gly Cys Asn Ala Leu Asn Leu Leu Val Gln Gly Ala Gly Gln Val Val  
 305 310 315 320

Ser Val Asp Cys Asn Pro Ala Gln Ser Ala Leu Leu Glu Leu Lys Lys  
 325 330 335  
 Val Ala Ile Gln Gln Leu Glu Phe Glu Asp Val Trp Gln Leu Phe Gly  
 340 345 350  
 Glu Gly Val His Pro Arg Ile Glu Glu Leu Tyr Glu Lys Lys Leu Ala  
 355 360 365  
 Pro Phe Leu Ser Gln Thr Ser His Asn Phe Trp Ser Lys Arg Leu Trp  
 370 375 380  
 Tyr Phe Gln His Gly Leu Tyr Tyr Gln Gly Gly Met Gly Lys Leu Cys  
 385 390 395 400  
 Trp Val Leu Gln Cys Leu Ala Val Val Leu Gly Leu Gly Lys Thr Val  
 405 410 415  
 Lys Arg Leu Ala Asn Ala Pro Thr Met Glu Glu Gln Arg Arg Leu Trp  
 420 425 430  
 Asp Ser Asn Met Leu Ile His Phe Val Lys Asn Gly Pro Lys Pro Leu  
 435 440 445  
 Val Trp Leu Phe Val Lys Phe Val Ser Leu Val Leu Phe Asn Lys Ala  
 450 455 460  
 Val Leu Trp Phe Gly Gly Gly Val Pro Gly Lys Gln Tyr Ala Leu Ile  
 465 470 475 480  
 Lys Ala Asp Gly Ile Pro Ile Glu Asn Tyr Ile Ala Arg Thr Met Asp  
 485 490 495  
 Gly Val Ala Glu Asn Ser His Val Arg Lys Gln Asn Tyr Phe Tyr Tyr  
 500 505 510  
 Asn Cys Leu Thr Gly Lys Phe Leu Arg Asp Asn Cys Pro Thr Tyr Leu  
 515 520 525  
 Arg Glu Ala Ala Phe Ala Thr Leu Lys Ser Gly Val Val Asp Asn Leu  
 530 535 540  
 Thr Val Ser Thr Asn Phe Phe Met Glu Glu Leu Lys Ala Arg Thr Tyr  
 545 550 555 560

Thr Lys Val Ile Leu Met Asp His Val Asp Trp Leu Asp Met Pro Val  
565 570 575

Ala Asn Glu Leu Ala Glu Cys Leu Ala Lys Gln Val Ala Pro Gly Gly  
580 585 590

Ile Val Ile Trp Arg Ser Ala Ser Leu Ser Pro Pro Tyr Ala Glu Leu  
595 600 605

Ile Gln Lys Ala Gly Phe Asp Val Arg Cys Ile Arg Arg Ala Thr Gln  
610 615 620

Gly Tyr Met Asp Arg Val Asn Met Tyr Ser Ser Phe Tyr Met Ala Arg  
625 630 635 640

Arg Lys Gly Ala Lys Lys Asp Asn  
645

<210> 46  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Synthetic

<400> 46  
caggatccaa tggggtcggg tcgt

24

<210> 47  
<211> 23  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Synthetic

<400> 47  
caggtaccgc cgccagctgc tta

23

<210> 48  
<211> 3427  
<212> DNA  
<213> Neurospora crassa

<400> 48  
crassagacc acgacgcaga atgacgaaca gagaaccaga catgtggatc acagaaccct 60  
gtctccctat gttactgggg gccgagcgct ggccacccat tccgttcgga tctcacgcaa 120  
aagcggagtg tggaggtggg gaacttcttt tacaacgctc gctattgaag cttctgaaga 180  
tgtacatacg ggtttatatg agagagaaaa gatataacta tcggtcttaa aggctaacct 240

caatccaact	tctttctctc	ttccttcggt	tcttgaacag	accgtcggat	cctctcttct	300
cttacttgac	accaacacct	tatagcaatg	ggagacaaca	gtgccatggc	ttctcatggg	360
ggtcacatgg	gtaacatttc	gctaccatca	ctttcagtaa	cccgaacact	agcagacctc	420
aacttcaaca	ccaccactac	caaatccatc	ttcttcaccg	gtgtggccgt	cttgggtattt	480
ctcgtcacca	ccagcaacta	ttcgcggaag	acaacaaaaa	acgaggacga	caatgaggat	540
gaaggcaacc	caagctccct	caaatccctc	ctcctcttct	gctactcctg	cttcatcaaa	600
cctcacgcca	ccgccggcac	cacaggaaca	cagcaagatg	ccctggagtc	gttttaccga	660
agtcaggcag	acatttatga	tgcgacaagg	ggtagcgtac	tgaaggggag	ggaggatatg	720
ttggctcttg	cagcttcgca	gttgaggtac	aaggttgagg	ctggacttgg	cggacttggg	780
ggagctggag	atgggcttga	aaaaagacaa	aggaatggga	agacatgtgt	aaccgtggcc	840
gggacagggg	ccgggacaag	gaggaaaccg	atatgggtag	atgtacgcct	cctcatcaga	900
acccttgcca	atgttgaaat	cccacacacg	gctgaccatc	gctacaaaac	aggctcgggtg	960
gggcacaggc	tggaatatcg	aagccatggc	caagtttgtc	aacgtctctg	aattcttcaa	1020
gactgtttac	ctagtggact	tttcgccgtc	actttgtgaa	gtggctagga	agcggtttgc	1080
caggctgggg	tgggagaatg	tgagagttaa	ctgcacggat	gctcgcaagt	ttaggcttga	1140
ggattatgag	gatgttgacg	aaggagagtc	tggctctgga	gattcttcgc	cttctttgtc	1200
gggttggtgg	ggggagacga	agccgggacg	acatgcggga	gctgagttga	tcaccatgtc	1260
ttatagcctt	tcgatgatgg	tatgtttctt	gattaaggac	gtcttggggg	ggtgtgtgct	1320
gacatatcat	gacgatgtag	ccggattatt	tctcgattat	cgattcgctc	gagtctctgt	1380
tagcacctca	cggcttgatt	gccgtcgtgg	acttttacgc	ccagtcgaaa	gtcgcacttca	1440
cattccgcaa	ctacacgggt	ggtcttatga	accgacacgt	tggtatattc	gcgcggaact	1500
tctggcgctc	gtggttcgat	gctgacaggg	tgtctcttga	gccagctcgt	cgagattatc	1560
tcgagtacag	gttcgggact	gtcctgaccg	tcaacgcccg	caacaacact	ttgggagcaa	1620
ttccttacta	catctggttg	ggatgcctca	agaagccctt	ttctacgtcg	agtctaccac	1680
acgaaattgt	ggaacacatc	gatgctattg	cgacagagtc	cccaagatca	tcaccccgtc	1740
tagtgggcaa	acattcttcc	tcagcaacaa	atgcgctagc	ctttgcagtc	ggccgcacag	1800
cgccggagat	gcgctcaaag	gccttcaata	cggccatcga	gaacatctcg	gccaacctac	1860
ctctcccgtc	cttcttctac	caaaatcacc	actggaggat	ctactacgac	gatcaactcc	1920
cgaagcacac	ccagttcaat	gacgagtaca	tctacgcctt	tacctgggaa	gactcgcgcg	1980
tcgacagaga	actccttaac	ctcgggcccc	acgacgtcgt	cctagccatc	accagcgccg	2040
gcgacaacat	tctttctctc	ctgatgcaga	gtcccgcctg	cgtgcacgcc	atcgacctaa	2100

accagccca aaaccacctg cttgaactca aagtcgcctc ttttacgact ctggattacc	2160
ccgacgtctg gaagatcttc ggtgagggca aacaccccgga ctttcgctca ctgctcatct	2220
ccaaactctc cctcacctc tccggccgcg cgtttcaata ctggctatcc aatgcgcaca	2280
tatttaccga ccctgcgggg cgcggtctct atgataccgg cggctcccga tacgctatcc	2340
gtttcttcg ctggatttcc acactcttct tctgccgctc cgcggtccgt cgacttctct	2400
ctactccac cctcgaagg caacgttcca tctaccacac caagattcgt ccctgtctgc	2460
tcaaccgctt cgtcaacggc ctggtcctca gctccgacgc ctctctctgg tcggctttgg	2520
gcgtgcccaa gaatcaagt gctatgatcg aagccgacta ccaccgccgt tctatctcct	2580
cctccaccac cccagcagc aaagaaaaac caagccgcgc cgaagcaatc ctccactaca	2640
caacctccac ccttgatccc gttctctcca cctccacct tgcctcggac aaccttact	2700
acctcgtctg tgtcctggga caatacacac gccagtgcc tcccgattac ctttcccctg	2760
ccgcccactc tatactcagc gctcctggag cctttgacgg cttacgcac cacacggatg	2820
aaatacagga ggtgttggt aggtttcagc cgggtacttt gacagtagcg gtggtgatgg	2880
atagtatgga ttggttcgat ccgccttcgc ctgaggagga aaaggaagga aggggcaagg	2940
cgagggagca agtgaggagg ttgaatcggg cgttgaaggt gggtggaag gtgttggtga	3000
ggagcgcggg agtgagaccg tggatgtga gggtttttgt ggaggaaggg tttggagcaa	3060
ggaggggtggg ttgtcgtgaa tccggaaggg gggaccagga gtgtattgac agggatgaata	3120
tgtatgctag ttgttgatc ttggaaaaga tggaggatct tgaggagtgt gttgactcag	3180
cttaagtaga tattttggtc ctaggtgata agttggataa ccctatttag cgcaagacta	3240
cctctactgg ccagaacacg gcctctgata ttctttatct gccactcaac cattcagccg	3300
gctttttgat tcaagcatgg taaaacgagg accgtgggag aaatcgcaac ctctttactg	3360
gcctgacatg acaaggcggg tcagataagt atttgttata agagcttcca cagcaaaaag	3420
atgtccc	3427

<210> 49  
 <211> 2727  
 <212> DNA  
 <213> *Neurospora crassa*

<400> 49	
atgggagaca acagtgccat ggcttctcat ggtggtcaca tgggtaacat ttcgctacca	60
tcactttcag taaccgaac actagcagac ctcaacttca acaccaccac taccaaatcc	120
atcttcttca ccggtgtggc cgtcttggtg tttctcgtca ccaccagcaa ctattcgcgg	180
aagacaacca aaaacgagga cgacaatgag gatgaaggca acccaagctc cctcaaattc	240
ctctctctct tctgctactc ctgcttcac aaacctcag ccaccgccgg caccacagga	300



acacagcaag	atgccctgga	gtcgttttac	cgaagtcagg	cagacattta	tgatgcgaca	360
aggggtacgc	tactgaaggg	gagggaggat	atgttggctc	ttgcagcttc	gcagttgagg	420
tacaaggttg	aggctggact	tggcggactt	ggaggagctg	gagatgggct	tgaaaaaaga	480
caaaggaatg	ggaagacatg	tgtaaccgtg	gccgggacag	ggaccgggac	aaggaggaaa	540
ccgatatggg	tagatgtcgg	tgggggcaca	ggctggaata	tcgaagccat	ggccaagt	600
gtcaacgtct	ctgaattctt	caagactgtt	tacctagtgg	acttttcgcc	gtcactttgt	660
gaagtggcta	ggaagcgggt	tgccaggctg	gggtgggaga	atgtgagagt	tatctgcacg	720
gatgctcgca	agtttaggct	tgaggattat	gaggatgttg	acgaaggaga	gtctggctct	780
ggagattctt	cgccttcttt	gtcgggttgg	tggggggaga	cgaagccggg	acgacatgcg	840
ggagctgagt	tgatcaccat	gtcttatagc	ctttcgatga	tgccggatta	tttctcgatt	900
atcgattcgc	tcgagtctct	gttagcacct	cacggcctga	ttgccgtcgt	ggacttttac	960
gcccagtcga	aagtcgactt	cacattccgc	aactacacgg	gtggtcttat	gaaccgacac	1020
gttggtatt	tcgcgcggaa	cttctggcgc	tcgtggttcg	atgctgacag	ggtgtctctt	1080
gagccagctc	gtcgagatta	tctcgagtac	aggttcggga	ctgtcctgac	cgtcaacgcc	1140
cgcaacaaca	ctttgggagc	aattccttac	tacatctggt	tgggatgcct	caagaagccc	1200
ttttctacgt	cgagtctacc	acacgaaatt	gtggaacaca	tcgatgctat	tgcgacagag	1260
tccccaagat	catcaccccg	tctagtgggc	aaacattctt	cctcagcaac	aaatgcgcta	1320
gcctttgcag	tcggccgcac	agcgcgggag	atgcgctcaa	aggccttcaa	tacggccatc	1380
gagaacatct	cggccaacct	acctctcccg	tccttcttct	acaaaaatca	ccactggagg	1440
atctactacg	acgatcaact	cccgaagcac	accagttca	atgacgagta	catctacgcc	1500
tttacctggg	aagactcgcg	cgtcgacaga	gaactcctta	acctcggggc	cgacgacgtc	1560
gtcctagcca	tcaccagcgc	cggcgacaac	attctttcct	acctgatgca	gagtcccgt	1620
cgcgtgcacg	ccatcgacct	aaaccagcc	caaaaccacc	tgcttgaact	caaagtcgcc	1680
tcttttacga	ctctggatta	ccccgacgtc	tggaagatct	tcggtgaggg	caaacacccc	1740
gactttcgct	cactgctcat	ctccaaactc	ttccctcacc	tctccggccg	cgcgtttcaa	1800
tactggctat	ccaatgcgca	catatttacc	gaccctgcgg	ggcgcggtct	ctatgatacc	1860
ggcggctccc	gatacgctat	ccgtttcttc	cgctggat	ccacactctt	cttctgccgc	1920
tccgcgggtcc	gtcgacttct	ctctactccc	accctcgaag	ggcaacgttc	catctaccac	1980
accaagattc	gtccctgtct	gtcgaaccgc	ttcgtcaacg	gcctggctct	cagctccgac	2040
gccttcctct	ggtcggcttt	gggcgtgccc	aagaatcaag	tggtatgat	cgaagccgac	2100
taccaccgcc	gttctatctc	ctcctccacc	accccagca	gcaaagaaaa	accaagccgc	2160

gccgaagcaa tcctccacta cacaacctcc acccttgatc ccgttctctc cacctcccac 2220  
 cttgcctcgg acaacctta ctacctcgtc tgtgtcctgg gacaatacac acgccagtgc 2280  
 catcccgatt acctttcccc tgccgcccac tctatactca gcgctcctgg agcctttgac 2340  
 ggcttacgca tccacacgga tgaaatacag gaggtgttgg ctaggtttca gccgggtact 2400  
 ttgacagtag cgggtggtgat ggatagtatg gattggttcg atccgccttc gcctgaggag 2460  
 gaaaaggaag gaaggggcaa ggcgaggag caagtgagga ggttgaatcg ggcgttgaag 2520  
 gtgggtggaa aggtgttggt gaggagcgcg ggagtggagc cgtggtatgt gaggggtttt 2580  
 gtggaggaag ggtttggagc aaggagggtg ggtgtcgtg aatccggaag gggggaccag 2640  
 gagtgtattg acagggtgaa tatgtatgct agttgttga tcttggaata gatggaggat 2700  
 cttgaggagt tgggtgactc agcttaa 2727

<210> 50  
 <211> 908  
 <212> PRT  
 <213> Neurospora crassa

<400> 50

Met Gly Asp Asn Ser Ala Met Ala Ser His Gly Gly His Met Gly Asn  
 1 5 10 15

Ile Ser Leu Pro Ser Leu Ser Val Thr Arg Thr Leu Ala Asp Leu Asn  
 20 25 30

Phe Asn Thr Thr Thr Thr Lys Ser Ile Phe Phe Thr Gly Val Ala Val  
 35 40 45

Leu Val Phe Leu Val Thr Thr Ser Asn Tyr Ser Arg Lys Thr Thr Lys  
 50 55 60

Asn Glu Asp Asp Asn Glu Asp Glu Gly Asn Pro Ser Ser Leu Lys Ser  
 65 70 75 80

Leu Leu Leu Phe Cys Tyr Ser Cys Phe Ile Lys Pro His Ala Thr Ala  
 85 90 95

Gly Thr Thr Gly Thr Gln Gln Asp Ala Leu Glu Ser Phe Tyr Arg Ser  
 100 105 110

Gln Ala Asp Ile Tyr Asp Ala Thr Arg Gly Thr Leu Leu Lys Gly Arg  
 115 120 125

Glu Asp Met Leu Ala Leu Ala Ala Ser Gln Leu Arg Tyr Lys Val Glu  
 130 135 140

Ala Gly Leu Gly Gly Leu Gly Gly Ala Gly Asp Gly Leu Glu Lys Arg  
 145 150 155 160  
 Gln Arg Asn Gly Lys Thr Cys Val Thr Val Ala Gly Thr Gly Thr Gly  
 165 170 175  
 Thr Arg Arg Lys Pro Ile Trp Val Asp Val Gly Gly Gly Thr Gly Trp  
 180 185 190  
 Asn Ile Glu Ala Met Ala Lys Phe Val Asn Val Ser Glu Phe Phe Lys  
 195 200 205  
 Thr Val Tyr Leu Val Asp Phe Ser Pro Ser Leu Cys Glu Val Ala Arg  
 210 215 220  
 Lys Arg Phe Ala Arg Leu Gly Trp Glu Asn Val Arg Val Ile Cys Thr  
 225 230 235 240  
 Asp Ala Arg Lys Phe Arg Leu Glu Asp Tyr Glu Asp Val Asp Glu Gly  
 245 250 255  
 Glu Ser Gly Ser Gly Asp Ser Ser Pro Ser Leu Ser Gly Trp Trp Gly  
 260 265 270  
 Glu Thr Lys Pro Gly Arg His Ala Gly Ala Glu Leu Ile Thr Met Ser  
 275 280 285  
 Tyr Ser Leu Ser Met Met Pro Asp Tyr Phe Ser Ile Ile Asp Ser Leu  
 290 295 300  
 Glu Ser Leu Leu Ala Pro His Gly Leu Ile Ala Val Val Asp Phe Tyr  
 305 310 315 320  
 Ala Gln Ser Lys Val Asp Phe Thr Phe Arg Asn Tyr Thr Gly Gly Leu  
 325 330 335  
 Met Asn Arg His Val Gly Tyr Phe Ala Arg Asn Phe Trp Arg Ser Trp  
 340 345 350  
 Phe Asp Ala Asp Arg Val Ser Leu Glu Pro Ala Arg Arg Asp Tyr Leu  
 355 360 365  
 Glu Tyr Arg Phe Gly Thr Val Leu Thr Val Asn Ala Arg Asn Asn Thr  
 370 375 380

Leu Gly Ala Ile Pro Tyr Tyr Ile Trp Leu Gly Cys Leu Lys Lys Pro  
 385 390 395 400

Phe Ser Thr Ser Ser Leu Pro His Glu Ile Val Glu His Ile Asp Ala  
 405 410 415

Ile Ala Thr Glu Ser Pro Arg Ser Ser Pro Arg Leu Val Gly Lys His  
 420 425 430

Ser Ser Ser Ala Thr Asn Ala Leu Ala Phe Ala Val Gly Arg Thr Ala  
 435 440 445

Pro Glu Met Arg Ser Lys Ala Phe Asn Thr Ala Ile Glu Asn Ile Ser  
 450 455 460

Ala Asn Leu Pro Leu Pro Ser Phe Phe Tyr Gln Asn His His Trp Arg  
 465 470 475 480

Ile Tyr Tyr Asp Asp Gln Leu Pro Lys His Thr Gln Phe Asn Asp Glu  
 485 490 495

Tyr Ile Tyr Ala Phe Thr Trp Glu Asp Ser Arg Val Asp Arg Glu Leu  
 500 505 510

Leu Asn Leu Gly Pro Asp Asp Val Val Leu Ala Ile Thr Ser Ala Gly  
 515 520 525

Asp Asn Ile Leu Ser Tyr Leu Met Gln Ser Pro Ala Arg Val His Ala  
 530 535 540

Ile Asp Leu Asn Pro Ala Gln Asn His Leu Leu Glu Leu Lys Val Ala  
 545 550 555 560

Ser Phe Thr Thr Leu Asp Tyr Pro Asp Val Trp Lys Ile Phe Gly Glu  
 565 570 575

Gly Lys His Pro Asp Phe Arg Ser Leu Leu Ile Ser Lys Leu Ser Pro  
 580 585 590

His Leu Ser Gly Arg Ala Phe Gln Tyr Trp Leu Ser Asn Ala His Ile  
 595 600 605

Phe Thr Asp Pro Ala Gly Arg Gly Leu Tyr Asp Thr Gly Gly Ser Arg  
 610 615 620

Tyr Ala Ile Arg Phe Phe Arg Trp Ile Ser Thr Leu Phe Phe Cys Arg  
 625 630 635 640

Ser Ala Val Arg Arg Leu Leu Ser Thr Pro Thr Leu Glu Gly Gln Arg  
 645 650 655

Ser Ile Tyr His Thr Lys Ile Arg Pro Cys Leu Leu Asn Arg Phe Val  
 660 665 670

Asn Gly Leu Val Leu Ser Ser Asp Ala Phe Leu Trp Ser Ala Leu Gly  
 675 680 685

Val Pro Lys Asn Gln Val Ala Met Ile Glu Ala Asp Tyr His Arg Arg  
 690 695 700

Ser Ile Ser Ser Ser Thr Thr Pro Ser Ser Lys Glu Lys Pro Ser Arg  
 705 710 715 720

Ala Glu Ala Ile Leu His Tyr Thr Thr Ser Thr Leu Asp Pro Val Leu  
 725 730 735

Ser Thr Ser His Leu Ala Ser Asp Asn Pro Tyr Tyr Leu Val Cys Val  
 740 745 750

Leu Gly Gln Tyr Thr Arg Gln Cys His Pro Asp Tyr Leu Ser Pro Ala  
 755 760 765

Ala His Ser Ile Leu Ser Ala Pro Gly Ala Phe Asp Gly Leu Arg Ile  
 770 775 780

His Thr Asp Glu Ile Gln Glu Val Leu Ala Arg Phe Gln Pro Gly Thr  
 785 790 795 800

Leu Thr Val Ala Val Val Met Asp Ser Met Asp Trp Phe Asp Pro Pro  
 805 810 815

Ser Pro Glu Glu Glu Lys Glu Gly Arg Gly Lys Ala Arg Glu Gln Val  
 820 825 830

Arg Arg Leu Asn Arg Ala Leu Lys Val Gly Gly Lys Val Leu Leu Arg  
 835 840 845

Ser Ala Gly Val Glu Pro Trp Tyr Val Arg Val Phe Val Glu Glu Gly  
 850 855 860

Phe Gly Ala Arg Arg Val Gly Cys Arg Glu Ser Gly Arg Gly Asp Gln  
 865 870 875 880

Glu Cys Ile Asp Arg Val Asn Met Tyr Ala Ser Cys Trp Ile Leu Glu  
 885 890 895

Lys Met Glu Asp Leu Glu Glu Leu Val Asp Ser Ala  
 900 905

<210> 51  
 <211> 33  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic

<400> 51  
 caggtaccgg atccaatagc aatgggagac aac

33

<210> 52  
 <211> 34  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic

<400> 52  
 caaagctttc tagactactt aagctgagtc aacc

34